Borwein–Preiss Variational Principle Revisited

A. Y. Kruger[a,*], S. Plubtieng[b], T. Seangwattana[b]

[a]*Centre for Informatics and Applied Optimization, Faculty of Science and Technology, Federation University Australia, Ballarat, Victoria 3353, Australia*
[b]*Department of Mathematics, Faculty of Science, Naresuan University, Phitsanulok 65000, Thailand*

Abstract

In this article, we refine and slightly strengthen the metric space version of the Borwein–Preiss variational principle due to Li, Shi, *J. Math. Anal. Appl.* 246, 308–319 (2000), clarify the assumptions and conclusions of their Theorem 1 as well as Theorem 2.5.2 in Borwein, Zhu, *Techniques of Variational Analysis*, Springer (2005) and streamline the proofs. Our main result, Theorem 3 is formulated in the metric space setting. When reduced to Banach spaces (Corollary 9), it extends and strengthens the smooth variational principle established in Borwein, Preiss, *Trans. Amer. Math. Soc.* 303, 517-527 (1987) along several directions.

Keywords: Borwein-Preiss variational principle, smooth variational principle, gauge-type function, perturbation

1. Introduction

The celebrated *Ekeland variational principle* [1] has been around for more than four decades. It almost immediately became one of the main tools in optimization theory and various branches of analysis. The number of publications containing "Ekeland variational principle" in their title has exceeded 200. Several other variational principles followed: due to Stegall [2], Borwein–Preiss [3], Deville–Godefroy–Zizler [4] and others.

---

*Corresponding author
Email addresses: `a.kruger@federation.edu.au` (A. Y. Kruger),
`somyotp@nu.ac.th` (S. Plubtieng), `seangwattana_t@hotmail.com` (T. Seangwattana)

Given an "almost minimal" point of a function, a variational principle guaranties the existence of another point and a suitably perturbed function for which this point is (strictly) minimal and provides estimates of the (generalized) distance between the points and also the size of the perturbation. Typically variational principles assume the underlying space to be complete metric (quasi-metric) or Banach and the function (sometimes vector- or set-valued) to possess a kind of semicontinuity.

The principles differ mainly in terms of the class of perturbations they allow. The perturbation guaranteed by the original Ekeland variational principle (valid in general complete metric spaces) is nonsmooth even if the underlying space is a smooth Banach space and the function is everywhere Fréchet differentiable. In contrast, the *Borwein–Preiss variational principle* (originally formulated in the Banach space setting) works with a special class of perturbations determined by the norm; when the space is *smooth* (i.e., the norm is Fréchet differentiable away from the origin), the perturbations are smooth too. Because of that, the Borwein–Preiss variational principle is referred to in [3] as the *smooth variational principle*. It has found numerous applications and paved the way for a number of other smooth principles including the one due to Deville–Godefroy–Zizler [4].

The statement of the next theorem mostly follows that of [5, Theorem 2.5.3].

Theorem 1 (Borwein–Preiss variational principle). *Let $(X, \|\cdot\|)$ be a Banach space and function $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Suppose that $\epsilon > 0$, $\lambda > 0$ and $p \geq 1$. If $x_0 \in X$ satisfies*

$$f(x_0) < \inf_X f + \epsilon, \tag{1}$$

then there exist a point $\bar{x} \in X$ and sequences $\{x_i\}_{i=1}^{\infty} \subset X$ and $\{\delta_i\}_{i=0}^{\infty} \subset \mathbb{R}_+ \setminus \{0\}$ such that $x_i \to \bar{x}$ as $i \to \infty$, $\sum_{i=0}^{\infty} \delta_i = 1$, and

(i) $\|\bar{x} - x_i\| \leq \lambda \ (i = 0, 1, \ldots)$;

(ii) $f(\bar{x}) + \dfrac{\epsilon}{\lambda^p} \sum_{i=0}^{\infty} \delta_i \|\bar{x} - x_i\|^p \leq f(x_0)$;

(iii) $f(x) + \dfrac{\epsilon}{\lambda^p} \sum_{i=0}^{\infty} \delta_i \|x - x_i\|^p > f(\bar{x}) + \dfrac{\epsilon}{\lambda^p} \sum_{i=0}^{\infty} \delta_i \|\bar{x} - x_i\|^p$ *for all $x \in X \setminus \{\bar{x}\}$.*

When X is a smooth space and $p > 1$, the perturbation functions involved in (ii) and (iii) of the above theorem are smooth.

Among the known extensions of the Borwein–Preiss variational principle, we mention the work by Li and Shi [6, Theorem 1], where the principle was extended to metric spaces (of course at the expense of losing the smoothness) by replacing $\|\cdot\|^p$ in (ii) and (iii) by a more general "gauge-type" function $\rho :$ $X \times X \to \mathbb{R}$. They also strengthened Theorem 1 by showing the existence of $\bar{x}$ and $\{x_i\}_{i=1}^\infty$ validating the appropriately adjusted conclusions of the theorem for any sequence $\{\delta_i\}_{i=0}^\infty \subset \mathbb{R}_+$ with $\delta_0 > 0$. This last advancement allowed the authors to cover the Ekeland variational principle which corresponds to setting $\delta_i = 0$ for $i = 1, 2, \ldots$ The result by Li and Shi was later adapted in Theorem 2.5.2 in the book by Borwein and Zhu [5].

Another important advancement was made by Loewen and Wang [7, Theorem 2.2] who constructed in the Banach space setting a special class of perturbations subsuming those used in Theorem 1 and established strong minimality in the analogue of the condition (iii) above; cf. [7, Definition 2.1]. Bednarczuk and Zagrodny [8] extended recently the Borwein–Preiss variational principle to vector-valued functions.

In this article which follows the ideas of [3, 6, 5], we refine and slightly strengthen the metric space version of the Borwein–Preiss variational principle due to Li and Shi [6], clarify the assumptions and conclusions of [6, Theorem 1] and [5, Theorem 2.5.2] and streamline the proofs. When reduced to Banach spaces (Corollary 9), our main result extends and strengthens Theorem 1 along several directions.

1) The assumption $p \geq 1$ for the power index in (ii) and (iii) is relaxed to just $p > 0$. Of course, if $p < 1$, then the perturbation function involved in (ii) and (iii) is not convex.

2) The strict inequality (1) is replaced by the corresponding nonstrict one:
$$f(x_0) \leq \inf_X f + \epsilon. \tag{2}$$

Note that δ_0 must satisfy
$$\delta_0 \geq (f(x_0) - \inf_X f)/\epsilon \tag{3}$$

(see Corollary 9). Hence, when $f(x_0) = \inf_X f + \epsilon$, one has $\delta_0 \geq 1$ and cannot ensure the equality $\sum_{i=0}^\infty \delta_i = 1$.

3) Instead of assuming the existence of $\bar{x}$, $\{x_i\}_{i=1}^\infty$ and $\{\delta_i\}_{i=0}^\infty$ with $\sum_{i=0}^\infty \delta_i = 1$ as in Theorem 1, we show that $\bar{x}$ and $\{x_i\}_{i=1}^\infty$ exist for any sequence $\{\delta_i\}_{i=0}^\infty \subset \mathbb{R}_+$ (the fact first observed in [6]) with δ_0 satisfying (3). The

3

latter restriction still leaves one enough freedom to choose positive numbers δ_i $(i = 1, 2, \ldots)$ such that $\sum_{i=0}^{\infty} \delta_i < \infty$, thus ensuring the convergence of the series involved in the left-hand side of condition (iii). In the case of the strict inequality (1), one can obviously satisfy that restriction with some $\delta_0 < 1$ and choose positive numbers δ_i $(i = 1, 2, \ldots)$ such that $\sum_{i=0}^{\infty} \delta_i = 1$.

4) Similarly to [6], conditions (ii) and (iii) in Theorem 1 are complemented by a pair of conditions which correspond to the case when only finitely many elements of the sequence δ_i $(i = 0, 1, \ldots)$ are nonzero. These conditions strengthen the corresponding conditions in [6].

5) The case when the series $\sum_{i=0}^{\infty} \delta_i$ is divergent is not excluded. We show that the series involved in condition (ii) (and the right-hand side of condition (iii)) is still convergent. However, the series in the left-hand side of condition (iii) can be divergent for some $x \in X$.

6) The inequalities in (i) can be replaced by $\|\bar{x} - x_0\| \leq \lambda$ and $\|\bar{x} - x_i\| \leq \epsilon_i$ $(i = 1, 2, \ldots)$, where $\{\epsilon_i\}_{i=1}^{\infty}$ is an arbitrary sequence of positive numbers.

The rest of the article is subdivided into three sections. In the next one, we present and prove our main result extending the Borwein–Preiss variational principle in metric spaces. Section 3 contains some discussions of the main result and provides several corollaries. In the final Section 4, we identify developing a "smooth" regularity theory as a possible application of the extended Borwein–Preiss variational principle.

Our basic notation is standard, cf. [9, 5, 10]. X stands for either a metric or a Banach space. A metric or a norm in X are denoted by $d(\cdot, \cdot)$ or $\| \cdot \|$, respectively. $\mathbb{N}$ denotes the set of all positive integers.

2. Extended Borwein–Preiss Variational Principle

In this section, we extend the metric space version of the Borwein–Preiss variational principle due to Li and Shi [6] (cf. [5]) which subsumes also the Ekeland variational principle.

The theorem below involves sequences indexed by $i \in \mathbb{N}$. The set of all indices is subdivided into two groups: with $i < N$ and $i \geq N$ where N is an 'integer' which is allowed to be infinite: $N \in \mathbb{N} \cup \{+\infty\}$. If $N = +\infty$, then the first subset of indices is infinite, while the second one is empty. This trick allows us to treat the cases of a finite and infinite set of indices within the same framework. Another convention in this section concerns summation over an empty set of indices: $\sum_{k=0}^{-1} a_k = 0$.

4

Following [6, Theorem 1] and [5, Definition 2.5.1], we are going to employ in the rest of the article the following concept of *gauge-type* function.

Definition 2. Let (X, d) be a metric space. We say that a continuous function $\rho : X \times X \to [0, \infty]$ is a gauge-type function if

(i) $\rho(x, x) = 0$ for all $x \in X$,
(ii) for any $\epsilon > 0$ there exists $\delta > 0$ such that, for all $y, z \in X$, inequality $\rho(y, z) \leq \delta$ implies $d(y, z) < \epsilon$.

Here comes the main result.

Theorem 3 (Extended Borwein–Preiss variational principle). *Let X be a complete metric space and a function $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Suppose that ρ is a gauge-type function, $\epsilon > 0$, $\{\epsilon_i\}_{i=1}^{\infty}$ and $\{\delta_i\}_{i=0}^{\infty}$ are sequences such that*

- $\epsilon_i > 0$ *for all* $i \in \mathbb{N}$ *and* $\epsilon_i \downarrow 0$ *as* $i \to \infty$;

- $\delta_i > 0$ *for all* $i < N$ *and* $\delta_i = 0$ *for all* $i \geq N$, *where* $N \in \mathbb{N} \cup \{+\infty\}$.

If $x_0 \in X$ satisfies (2), then there exist a point $\bar{x} \in X$ and a sequence $\{x_i\}_{i=1}^{\infty} \subset X$ such that $x_i \to \bar{x}$ as $i \to \infty$ and

(i) $\rho(\bar{x}, x_0) \leq \epsilon/\delta_0$;
(ii) $\rho(\bar{x}, x_i) \leq \epsilon_i$ $(i = 1, 2, \ldots)$;
(iii) *if $N = +\infty$, then the series $\sum_{i=0}^{\infty} \delta_i \rho(\bar{x}, x_i)$ is convergent and*

$$f(\bar{x}) + \sum_{i=0}^{\infty} \delta_i \rho(\bar{x}, x_i) \leq f(x_0); \qquad (4)$$

otherwise the series $\sum_{i=N-1}^{\infty} \rho(x_{i+1}, x_i)$ is convergent and

$$f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i)$$
$$+ \delta_{N-1} \sup_{n \geq N-1} \left(\sum_{i=N-1}^{n-1} \rho(x_{i+1}, x_i) + \rho(\bar{x}, x_n) \right) \leq f(x_0); \quad (5)$$

(iv) *if $N = +\infty$, then*

$$f(x) + \sum_{i=0}^{\infty} \delta_i \rho(x, x_i) > f(\bar{x}) + \sum_{i=0}^{\infty} \delta_i \rho(\bar{x}, x_i) \quad \text{for all} \quad x \in X \setminus \{\bar{x}\}; \quad (6)$$

otherwise, for any $x \in X \setminus \{\bar{x}\}$, there exists an $m_0 \geq N$ such that, for all $m \geq m_0$,

$$f(x) + \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, x_m) > f(\bar{x})$$

$$+ \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) + \delta_{N-1} \sup_{n \geq m} \left(\sum_{i=m}^{n-1} \rho(x_{i+1}, x_i) + \rho(\bar{x}, x_n) \right). \quad (7)$$

Proof. (i) and (ii) We define sequences $\{x_i\}$ and $\{S_i\}$ inductively. Set

$$S_0 := \{x \in X \mid f(x) + \delta_0 \rho(x, x_0) \leq f(x_0)\}. \quad (8)$$

Obviously, $x_0 \in S_0$. Since the function $x \to f(x) + \delta_0 \rho(x, x_0)$ is lower semi-continuous, subset S_0 is closed. For any $x \in S_0$, we have

$$\rho(x, x_0) \leq \frac{f(x_0) - f(x)}{\delta_0} \leq \frac{\epsilon}{\delta_0}. \quad (9)$$

For $i = 0, 1, \ldots$, denote $j_i := \min\{i, N - 1\}$, i.e., j_i is the largest integer $j \leq i$ such that $\delta_j > 0$. Let $i \in \mathbb{N}$ and suppose $x_0, \ldots, x_{i-1}$ and $S_0, \ldots, S_{i-1}$ have been defined. We choose $x_i \in S_{i-1}$ such that

$$f(x_i) + \sum_{k=0}^{j_i-1} \delta_k \rho(x_i, x_k) \leq \inf_{x \in S_{i-1}} \left(f(x) + \sum_{k=0}^{j_i-1} \delta_k \rho(x, x_k) \right) + \delta_{j_i} \epsilon_i \quad (10)$$

and define

$$S_i := \left\{ x \in S_{i-1} \mid f(x) + \delta_{j_i} \rho(x, x_i) \right.$$

$$\left. + \sum_{k=0}^{j_i-1} \delta_k (\rho(x, x_k) - \rho(x_i, x_k)) \leq f(x_i) \right\}. \quad (11)$$

6

Obviously, $x_i \in S_i$. Since the function $x \to f(x) + \delta_{j_i}\rho(x, x_i) + \sum_{k=0}^{j_i-1} \delta_k\rho(x, x_k)$ is lower semicontinuous, subset S_i is closed. For any $x \in S_i$, we have

$$f(x) - f(x_i) + \sum_{k=0}^{j_i-1} \delta_k(\rho(x, x_k) - \rho(x_i, x_k)) + \delta_{j_i}\rho(x, x_i) \leq 0,$$

and consequently, making use of (10),

$$\rho(x, x_i) \leq \frac{1}{\delta_{j_i}}\left(f(x_i) + \sum_{k=0}^{j_i-1} \delta_k\rho(x_i, x_k) \right.$$
$$\left. - \left(f(x) + \sum_{k=0}^{j_i-1} \delta_k\rho(x, x_k) \right) \right) \leq \epsilon_i. \quad (12)$$

We can see that, for all $i \in \mathbb{N}$, subsets S_i are nonempty and closed, $S_i \subset S_{i-1}$, and $\sup_{x \in S_i} \rho(x, x_i) \to 0$ as $i \to \infty$. Since ρ is a gauge-type function, we also have $\sup_{x \in S_i} d(x, x_i) \to 0$ and consequently, $\mathrm{diam}(S_i) \to 0$. Since X is complete, $\cap_{i=0}^{\infty} S_i$ contains exactly one point; let it be $\bar{x}$. Hence, $\rho(\bar{x}, x_i) \to 0$ and $x_i \to \bar{x}$ as $i \to \infty$. Thanks to (9) and (12), $\bar{x}$ satisfies (i) and (ii).

Before proceeding to the proof of claim (iii), we prepare several building blocks which are going to be used when proving claims (iii) and (iv).

Let integers m, n and i satisfy $0 \leq m \leq i \leq n$. Since $x_{i+1} \in S_i$ and $\bar{x} \in S_n$, it follows from (8) (when $i = 0$) and (11) that

$$f(x_{i+1}) + \sum_{k=0}^{j_i-1} \delta_k(\rho(x_{i+1}, x_k) - \rho(x_i, x_k)) + \delta_{j_i}\rho(x_{i+1}, x_i) \leq f(x_i), \quad (13)$$

$$f(\bar{x}) + \sum_{k=0}^{j_n-1} \delta_k(\rho(\bar{x}, x_k) - \rho(x_n, x_k)) + \delta_{j_n}\rho(\bar{x}, x_n) \leq f(x_n). \quad (14)$$

We are going to add together inequalities (13) from $i = m$ to $i = n - 1$ and inequality (14). Depending on the value of N, three cases are possible.

If $N > n$, then $j_i = i$ and $j_n = n$. Adding inequalities (13) from $i = m$ to $i = n - 1$, we obtain

$$f(x_n) + \sum_{k=0}^{n-1} \delta_k\rho(x_n, x_k) - \sum_{k=0}^{m-1} \delta_k\rho(x_m, x_k) \leq f(x_m).$$

7

Adding the last inequality and inequality (14), we arrive at

$$f(\bar{x}) + \sum_{k=0}^{n} \delta_k \rho(\bar{x}, x_k) - \sum_{k=0}^{m-1} \delta_k \rho(x_m, x_k) \le f(x_m). \qquad (15)$$

If $N \le m$, then $j_i = N - 1$ and $j_n = N - 1$. Adding inequalities (13) from $i = m$ to $i = n - 1$, we obtain

$$f(x_n) + \sum_{k=0}^{N-2} \delta_k(\rho(x_n, x_k) - \rho(x_m, x_k)) + \delta_{N-1} \sum_{k=m}^{n-1} \rho(x_{k+1}, x_k) \le f(x_m).$$

Adding the last inequality and inequality (14), we arrive at

$$f(\bar{x}) + \sum_{k=0}^{N-2} \delta_k(\rho(\bar{x}, x_k) - \rho(x_m, x_k))$$
$$+ \delta_{N-1} \left(\sum_{k=m}^{n-1} \rho(x_{k+1}, x_k) + \rho(\bar{x}, x_n) \right) \le f(x_m). \qquad (16)$$

If $m < N \le n$, we add inequalities (13) separately from $i = m$ to $i = N-1$ and from $i = N$ to $i = n - 1$ and obtain, respectively,

$$f(x_N) + \sum_{k=0}^{N-1} \delta_k \rho(x_N, x_k) - \sum_{k=0}^{m-1} \delta_k \rho(x_m, x_k) \le f(x_m),$$

$$f(x_n) + \sum_{k=0}^{N-2} \delta_k(\rho(x_n, x_k) - \rho(x_N, x_k)) + \delta_{N-1} \sum_{k=N}^{n-1} \rho(x_{k+1}, x_k) \le f(x_N).$$

Adding the last two inequalities and inequality (14) together, we arrive at

$$f(\bar{x}) + \sum_{k=0}^{N-2} \delta_k \rho(\bar{x}, x_k) - \sum_{k=0}^{m-1} \delta_k \rho(x_m, x_k)$$
$$+ \delta_{N-1} \left(\sum_{k=N-1}^{n-1} \rho(x_{k+1}, x_k) + \rho(\bar{x}, x_n) \right) \le f(x_m). \qquad (17)$$

(iii) When $N = +\infty$, we set $m = 0$ in the inequality (15):

$$f(\bar{x}) + \sum_{k=0}^{n} \delta_k \rho(\bar{x}, x_k) \le f(x_0).$$

8

This inequality must hold for all $n \in \mathbb{N}$. Hence, the series $\sum_{k=0}^{\infty} \delta_k \rho(\bar{x}, x_k)$ is convergent and condition (4) holds true.

When $N < +\infty$, we set $m = 0$ and take $n = N - 1$ in the inequality (15) and any $n \geq N$ in the inequality (17):

$$f(\bar{x}) + \sum_{k=0}^{N-1} \delta_k \rho(\bar{x}, x_k) \leq f(x_0),$$

$$f(\bar{x}) + \sum_{k=0}^{N-2} \delta_k \rho(\bar{x}, x_k) + \delta_{N-1} \left(\sum_{i=N-1}^{n-1} \rho(x_{i+1}, x_i) + \rho(\bar{x}, x_n) \right) \leq f(x_0).$$

Since $\rho(\bar{x}, x_n) \to 0$ as $n \to \infty$, it follows from the last inequality that the series $\sum_{i=N-1}^{\infty} \rho(x_{i+1}, x_i)$ is convergent. Combining the two inequalities produces estimate (5).

(iv) For any $x \neq \bar{x}$, there exists an $m_0 \in \mathbb{N}$ such that $x \notin S_m$ for all $m \geq m_0$. By (11), this means that

$$f(x) + \sum_{k=0}^{j_m-1} \delta_k (\rho(x, x_k) - \rho(x_m, x_k)) + \delta_{j_m} \rho(x, x_m) > f(x_m). \qquad (18)$$

Depending on the value of N, we consider two cases.

<u>If $N = +\infty$</u>, then $j_m = m$. Since the series $\sum_{k=0}^{\infty} \delta_k \rho(\bar{x}, x_k)$ is convergent, we can pass in (15) to the limit as $n \to \infty$ to obtain

$$f(\bar{x}) + \sum_{k=0}^{\infty} \delta_k \rho(\bar{x}, x_k) \leq f(x_m) + \sum_{k=0}^{m-1} \delta_k \rho(x_m, x_k).$$

Subtracting the last inequality from (18), we arrive at

$$f(x) + \sum_{k=0}^{m} \delta_k \rho(x, x_k) > f(\bar{x}) + \sum_{k=0}^{\infty} \delta_k \rho(\bar{x}, x_k).$$

Condition (6) follows immediately.

<u>If $N < \infty$</u>, we can take $m_0 \geq N$. Then $j_m = N - 1$ and it follows from (16) that

$$f(\bar{x}) + \sum_{k=0}^{N-2} \delta_k (\rho(\bar{x}, x_k) - \rho(x_m, x_k))$$

$$+ \delta_{N-1} \sup_{n \geq m} \left(\sum_{k=m}^{n-1} \rho(x_{k+1}, x_k) + \rho(\bar{x}, x_n) \right) \leq f(x_m).$$

9

Subtracting the last inequality from (18), we arrive at (7). □

3. Comments and Corollaries

In this section, we discuss the main result proved in Section 2 and formulate a series of remarks and several corollaries.

Remark 4. 1. The series $\sum_{i=0}^{\infty} \delta_i \rho(x, x_i)$ in (6) does not have to be convergent for all $x \in X \setminus \{\bar{x}\}$.

2. If $N < \infty$, in the proof of part (iv) of Theorem 3 one can also consider the case $m_0 < N$. Then, for $m_0 \leq m < N$, one has $j_m = m$ and it follows from (17) that

$$f(\bar{x}) + \sum_{k=0}^{N-2} \delta_k \rho(\bar{x}, x_k) - \sum_{k=0}^{m-1} \delta_k \rho(x_m, x_k)$$
$$+ \delta_{N-1} \sup_{n \geq N} \left(\sum_{k=N-1}^{n-1} \rho(x_{k+1}, x_k) + \rho(\bar{x}, x_n) \right) \leq f(x_m).$$

Subtracting the last inequality from (18), one arrives at

$$f(x) + \sum_{i=0}^{m} \delta_i \rho(x, x_i) > f(\bar{x})$$
$$+ \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) + \delta_{N-1} \sup_{n \geq N} \left(\sum_{k=N-1}^{n-1} \rho(x_{k+1}, x_k) + \rho(\bar{x}, x_n) \right). \quad (19)$$

This estimate compliments (7).

3. Instead of ϵ-minimality in the sense of (2), it is sufficient to assume in Theorem 3 a weaker form of ϵ-minimality: $f(x) \geq f(x_0) - \epsilon$ for all $x \in X$ such that $f(x) + \delta_0 \rho(x, x_0) > f(x_0)$.

4. Looking at the statement of Theorem 3, it is easy to notice that considering a gauge-type function ρ and a sequence of positive numbers $\{\delta_i\}_{i=0}^{\infty}$ can be replaced by that of a sequence of gauge-type functions $\{\rho_i\}_{i=0}^{\infty}$ such that, for $i = 1, 2, \ldots$, function ρ_i is a multiple of ρ_0. The latter assumption can be relaxed or dropped at the expense of weakening or dropping the estimates in part (ii) of the concluding part of Theorem 3.

Moreover, one can modify the proof employing in it a sequence of functions $\{\rho_i\}_{i=0}^{\infty}$ which do not have to possess the second property in Definition 2,

as long as they ensure that the resulting sets S_i (cf. (11)) are closed and form a decreasing sequence with their diameters going to zero. This way one can establish additional properties of the sequence $\{x_i\}_{i=1}^{\infty}$ and its limiting point $\bar{x}$. An interesting example of such a sequence in a Banach space setting was considered by Loewen and Wang [7] who proved a strong variant of the Borwein–Preiss variational principle (with $\bar{x}$ being a strong minimizer of the corresponding perturbed function; cf. [7, Definition 2.1]).

5. Setting $\epsilon_i := \epsilon/(2^i \delta_0)$ $(i = 1, 2, \ldots)$, one can make the estimates in (ii) look as in [6, Theorem 1] and [5, Theorem 2.5.2].

6. Given a positive number λ, we can rewrite the conclusion of Theorem 3 in a more conventional form with $\delta_0 = 1$, $\rho(\bar{x}, x_0) \leq \lambda$ instead of (i) and conditions (4) and (6) replaced, respectively, with the following ones:

$$f(\bar{x}) + \frac{\epsilon}{\lambda} \sum_{i=0}^{\infty} \delta_i \rho(\bar{x}, x_i) \leq f(x_0), \tag{4'}$$

$$f(x) + \frac{\epsilon}{\lambda} \sum_{i=0}^{\infty} \delta_i \rho(x, x_i) > f(\bar{x}) + \frac{\epsilon}{\lambda} \sum_{i=0}^{\infty} \delta_i \rho(\bar{x}, x_i) \quad \text{for all} \quad x \in X \setminus \{\bar{x}\} \tag{6'}$$

and similar amendments in conditions (5), (7) and (19).

The next corollary gives some direct consequences of conditions (5) and (7) in Theorem 3.

Corollary 5. *Suppose all the assumptions of Theorem 3 are satisfied, and $N < \infty$. Then*

$$f(\bar{x}) + \sum_{i=0}^{N-1} \delta_i \rho(\bar{x}, x_i) \leq f(x_0), \tag{20}$$

$$f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) + \delta_{N-1} \sum_{i=N-1}^{\infty} \rho(x_{i+1}, x_i) \leq f(x_0), \tag{21}$$

11

and, for any $x \in X \setminus \{\bar{x}\}$, there exists an $m_0 \geq N$ such that, for all $m \geq m_0$,

$$f(x) + \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, x_m)$$

$$> f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) + \delta_{N-1} \rho(\bar{x}, x_m), \quad (22)$$

$$f(x) + \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, x_m)$$

$$> f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) + \delta_{N-1} \sum_{i=m}^{\infty} \rho(x_{i+1}, x_i), \quad (23)$$

and consequently,

$$f(x) + \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, \bar{x})$$

$$\geq f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i) \quad for \; all \quad x \in X, \quad (24)$$

where $\bar{x}$ and $\{x_i\}_{i=1}^{\infty}$ are a point and a sequence guaranteed by Theorem 3.

Proof. Conditions (20) and (21) correspond, respectively, to setting $n = N-1$ and letting $n \to \infty$ under the sup in condition (5). Similarly, conditions (22) and (23) correspond, respectively, to setting $n = m$ and letting $n \to \infty$ under the sup in condition (7). Condition (24) is obviously true when $x = \bar{x}$. When $x \neq \bar{x}$, it results from passing to the limit as $m \to \infty$ in any of the conditions (22) and (23) thanks to the continuity of ρ. □

Remark 6. 1. Conditions (20) and (21) are in general independent. Conditions (22) and (23) are independent too. Conditions (20) and (22) were formulated in [6]. Thanks to Corollary 5, Theorem 3 strengthens [6, Theorem 1].

2. In accordance with Theorem 3 and Corollary 5, $\bar{x}$ is a point of minimum of the sum $f + g$, where the perturbation function g is defined for $x \in X$ either as $g(x) := \sum_{i=0}^{\infty} \delta_i \rho(x, x_i)$ if $N = +\infty$ or as $g(x) := \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, \bar{x})$ otherwise. When $N = +\infty$, the minimum is strict. Thanks to the next proposition, if function ρ possesses the triangle inequality, the minimum is strict also when $N < +\infty$.

Recall that a function $\rho : X \times X \to \mathbb{R}$ possesses the *triangle inequality* if $\rho(x_1, x_3) \leq \rho(x_1, x_2) + \rho(x_2, x_3)$ for all $x_1, x_2, x_3 \in X$.

Proposition 7. *Along with conditions* (20)–(22), *consider the following one:*

$$f(x) + \sum_{i=0}^{N-2} \delta_i \rho(x, x_i) + \delta_{N-1} \rho(x, \bar{x}) > f(\bar{x}) + \sum_{i=0}^{N-2} \delta_i \rho(\bar{x}, x_i). \qquad (25)$$

If function ρ possesses the triangle inequality, then (21) $\Rightarrow$ (20) *and* (23) $\Rightarrow$ (22) $\Rightarrow$ (25).

Proof. For any $m, n \in \mathbb{N}$ with $m < n$, we have

$$\rho(\bar{x}, x_m) \leq \rho(\bar{x}, x_n) + \sum_{i=m}^{n-1} \rho(x_{i+1}, x_i),$$

and consequently, passing to the limit as $n \to \infty$,

$$\rho(\bar{x}, x_m) \leq \sum_{i=m}^{\infty} \rho(x_{i+1}, x_i).$$

Hence, (21) $\Rightarrow$ (20) and (23) $\Rightarrow$ (22). Condition (25) follows from (22) thanks to the inequality $\rho(x, x_m) \leq \rho(x, \bar{x}) + \rho(\bar{x}, x_m)$. $\qquad \square$

Corollary 8. *Suppose all the assumptions of Theorem 3 are satisfied, $N < +\infty$, and function ρ possesses the triangle inequality. Then condition* (25) *holds true for all $x \in X \setminus \{\bar{x}\}$.*

Proof. The statement is a consequence of Corollary 5 thanks to Proposition 7. $\qquad \square$

The next two statements are consequences of Theorem 3 when $N = +\infty$ and $N = 1$, respectively, and ρ is of a special form. The first one corresponds to the case $N = +\infty$, X a Banach space and $\rho(x_1, x_2) := \|x_1 - x_2\|^p$ where $p > 0$.

Corollary 9. *Let $(X, \|\cdot\|)$ be a Banach space and function $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Suppose that λ, p, ϵ, ϵ_i $(i = 1, 2, \ldots)$, δ_i $(i = 0, 1, \ldots)$ are positive numbers and $\epsilon_i \downarrow 0$ as $i \to \infty$. If $x_0 \in X$ and δ_0 satisfy conditions* (2) *and* (3), *then there exist a point $\bar{x} \in X$ and a sequence $\{x_i\}_{i=1}^{\infty} \subset X$ such that $x_i \to \bar{x}$ as $i \to \infty$ and*

13

(i) $\|\bar{x} - x_0\| \leq \lambda$;

(ii) $\|\bar{x} - x_i\| \leq \epsilon_i$ $(i = 1, 2, \ldots)$;

(iii) $f(\bar{x}) + \dfrac{\epsilon}{\lambda^p} \displaystyle\sum_{i=0}^{\infty} \delta_i \|\bar{x} - x_i\|^p \leq f(x_0)$;

(iv) $f(x) + \dfrac{\epsilon}{\lambda^p} \displaystyle\sum_{i=0}^{\infty} \delta_i \|x - x_i\|^p > f(\bar{x}) + \dfrac{\epsilon}{\lambda^p} \displaystyle\sum_{i=0}^{\infty} \delta_i \|\bar{x} - x_i\|^p$ for all $x \in X \setminus \{\bar{x}\}$.

Proof. Set $\rho(x_1, x_2) := \|x_1 - x_2\|^p$, $x_1, x_2 \in X$. It is easy to check that ρ is a gauge-type function. Set $\epsilon' := \epsilon \delta_0$, $\epsilon_i' := \epsilon_i^p$ $(i = 1, 2, \ldots)$, $\delta_i' := (\epsilon/\lambda^p)\delta_i$ $(i = 0, 1, \ldots)$. Then $f(x_0) \leq \inf_X f + \epsilon'$, $\epsilon_i' \downarrow 0$ as $i \to \infty$ and $\epsilon'/\delta_0' = \lambda^p$. The conclusion follows from Theorem 3 with ϵ', ϵ_i' and δ_i' in place of ϵ, ϵ_i and δ_i, respectively. $\qquad\square$

Condition (iv) means that $\bar{x}$ is a point of strict minimum of the function $x \mapsto f(x) + (\epsilon/\lambda^p)g(x)$, where $g(x) := \sum_{i=0}^{\infty} \delta_i \|x - x_i\|^p$. If X is Fréchet smooth, $p > 1$, and $\sum_{i=0}^{\infty} \delta_i < \infty$, then g is defined on the whole of X and is everywhere Fréchet differentiable, i.e., we have an example of a smooth variational principle of Borwein–Preiss type.

Remark 10. 1. Apart from (3), no other restrictions are imposed on the positive numbers δ_i, $i = 0, 1, \ldots$

2. Condition (2) does not exclude the equality case: $f(x_0) = \inf_X f + \epsilon$. In the latter case, condition (3) is equivalent to $\delta_0 \geq 1$. This still allows one to chose positive numbers δ_i, $i = 1, 2, \ldots$, such that $\sum_{i=0}^{\infty} \delta_i < \infty$ if necessary.

When the inequality (2) is strict, then one can choose $\delta_0 < 1$ and positive numbers δ_i, $i = 1, 2, \ldots$, such that $\sum_{i=0}^{\infty} \delta_i = 1$.

The next statement is the Ekeland variational principle. It corresponds to $N = 1$ and ρ being a distance function.

Corollary 11. *Let (X, d) be a complete metric space and function $f : X \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous. Suppose $\lambda > 0$ and $\epsilon > 0$. If $x_0 \in X$ satisfies (2), then there exists a point $\bar{x} \in X$ such that*

(i) $d(\bar{x}, x_0) \leq \lambda$;

(ii) $f(\bar{x}) + \dfrac{\epsilon}{\lambda} d(\bar{x}, x_0) \leq f(x_0)$;

(iii) $f(x) + \dfrac{\epsilon}{\lambda} d(x, \bar{x}) > f(\bar{x})$ *for all $x \in X \setminus \{\bar{x}\}$.*

Proof. Set $\rho := d$, $N = 1$, $\delta_0 := \epsilon/\lambda$, $\epsilon_i := \epsilon/2^i$ and $\delta_i := 0$ $(i = 1, 2, \ldots)$. Then $\epsilon_i \downarrow 0$ as $i \to \infty$ and $\epsilon/\delta_0 = \lambda$. The conclusion follows from Theorem 3 and Corollary 8. $\qquad\square$

4. "Smooth" Regularity Theory

One can try to use the estimates in Theorem 3 for developing a "smooth" regularity theory similar to the conventional theory based on the application of the Ekeland variational principle (cf. [11, 12, 10]) and usually using certain *slopes* to formulate primal space criteria (cf. [11, 12, 13, 14]). The first step towards the development of such a theory would be defining appropriate "smooth" slopes.

To illustrate the idea, we consider briefly the case $N = +\infty$. Let a function $f : X \to \mathbb{R} \cup \{+\infty\}$, a gauge-type function $\rho : X \times X \to [0, \infty]$ and a sequence $\{\delta_i\}_{i=0}^{\infty} \subset \mathbb{R}_+ \setminus \{0\}$ with $\delta_0 = 1$ be given.

For a sequence $\{x_i\}_{i=0}^{\infty} \subset X$, define

$$g_{\{x_i\}}(u) := \sum_{i=0}^{\infty} \delta_i \rho(u, x_i), \quad u \in X.$$

Next, for an $x \in X$ with $f(x) < \infty$ and a sequence $\{x_i\}_{i=0}^{\infty} \subset X$ convergent to x with $g_{\{x_i\}}(x) < \infty$, the slope of f at $(x, \{x_i\})$ can be defined as follows:

$$|\nabla f|(x, \{x_i\}) := \limsup_{\substack{u \to x \\ g_{\{x_i\}}(u) \neq g_{\{x_i\}}(x)}} \frac{[f(x) - f(u)]_+}{g_{\{x_i\}}(u) - g_{\{x_i\}}(x)}. \tag{26}$$

Similarly to the conventional slope, this quantity characterizes the maximal 'rate of descent' of f at x (with respect to $g_{\{x_i\}}$).

Theorem 3 implies the existence of a point $\bar{x} \in X$ near the given point x_0 and a sequence $\{x_i\}_{i=1}^{\infty} \subset X$ convergent to $\bar{x}$ such that $|\nabla f|(\bar{x}, \{x_i\})$ is small. Moreover, it provides quantitative estimates for $|\nabla f|(\bar{x}, \{x_i\})$ and the 'distance' (in terms of ρ) from $\bar{x}$ to x_0. More specifically, in the framework of Remark 4.6, one has $\rho(\bar{x}, x_0) \leq \lambda$ and $|\nabla f|(\bar{x}, \{x_i\}) \leq \epsilon/\lambda$.

Furthermore, since (6) (and (6′)) is a global condition, it could make sense to incorporate along with the slope (26) a *nonlocal* analogue of (26) as well as their *strict* (outer) extensions along the lines of [13, 14]. In the Banach space setting and with ρ appropriately defined (cf. Corollary 9), one can try to define a dual space counterpart of (26) and formulate subdifferential consequences of Theorem 3 exploiting the original idea of Borwein and Preiss [3].

This type of conditions should be useful when developing "smooth" criteria of error bounds and metric (Hölder) (sub-)regularity along the lines of [13, 14].

15

The case $N < \infty$ is also of interest and can be handled in a similar way. The appropriate definitions of slopes can be derived from condition (7) (or its 'm-free' consequence (24)).

This topic goes beyond the scope of the current article and is left for future research. Extending Theorem 3 and its corollaries to vector-valued functions seems to be another interesting direction of future research.

Acknowledgments

References

References

[1] I. Ekeland, On the variational principle, J. Math. Anal. Appl. 47 (1974) 324–353.

[2] C. Stegall, Optimization of functions on certain subsets of Banach spaces, Math. Ann. 236 (2) (1978) 171–176. `doi:10.1007/BF01351389`.

[3] J. M. Borwein, D. Preiss, A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions, Trans. Amer. Math. Soc. 303 (2) (1987) 517–527. `doi:10.2307/2000681`.

[4] R. Deville, G. Godefroy, V. Zizler, Smoothness and Renormings in Banach spaces, Vol. 64 of Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Scientific & Technical, Harlow, 1993.

[5] J. M. Borwein, Q. J. Zhu, Techniques of Variational Analysis, Springer, New York, 2005.

[6] Y. Li, S. Shi, A generalization of Ekeland's ϵ-variational principle and its Borwein-Preiss smooth variant, J. Math. Anal. Appl. 246 (1) (2000) 308–319.

[7] P. D. Loewen, X. Wang, A generalized variational principle, Canad. J. Math. 53 (6) (2001) 1174–1193. `doi:10.4153/CJM-2001-044-8`.

[8] E. M. Bednarczuk, D. Zagrodny, A smooth vector variational principle, SIAM J. Control Optim. 48 (6) (2010) 3735–3745. `doi:10.1137/090758271`.

[9] R. T. Rockafellar, R. J.-B. Wets, Variational Analysis, Springer, Berlin, 1998.

[10] A. L. Dontchev, R. T. Rockafellar, Implicit Functions and Solution Mappings. A View from Variational Analysis, 2nd Edition, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2014.

[11] D. Azé, A unified theory for metric regularity of multifunctions, J. Convex Anal. 13 (2) (2006) 225–252.

[12] A. D. Ioffe, Metric regularity and subdifferential calculus, Russian Math. Surveys 55 (2000) 501–558.

[13] A. Y. Kruger, Error bounds and metric subregularity, Optimization 64 (1) (2015) 49–79. `doi:10.1080/02331934.2014.938074`.

[14] A. Y. Kruger, Error bounds and Hölder metric subregularity, Set-Valued Var. Anal. ? (2015) 1–32. `doi:10.1007/s11228-015-0330-y`.